An association between variants in the IGF2 gene and Beckwith–Wiedemann syndrome: interaction between genotype and epigenotype

Adele Murrell1, Sarah Heeson1, Wendy N. Cooper2, Eleanor Douglas1, Sophia Apostolidou3, Gudrun E. Moore3, Eamonn R. Maher2 and Wolf Reik1,*

1Laboratory of Developmental Genetics and Imprinting, Developmental Genetics Programme, The Babraham Institute, Cambridge, UK, 2Section of Medical and Molecular Genetics, Department of Paediatrics and Child Health, The Medical School, The University of Birmingham, Birmingham, UK and 3Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College London, Hammersmith Campus, London, UK

Received October 13, 2003; Revised and Accepted November 8, 2003

Beckwith–Wiedemann syndrome (BWS) is a fetal overgrowth disorder involving the deregulation of a number of genes, including IGF2 and CDKN1C, in the imprinted gene cluster on chromosome 11p15.5. In sporadic BWS cases the majority of patients have epimutations in this region. Loss of imprinting of the IGF2 gene is frequently observed in BWS, as is reduced CDKN1C expression related to loss of maternal allele-specific methylation (LOM) of the differentially methylated region KvDMR1. The causes of epimutations are unknown, although recently an association with assisted reproductive technologies has been described. To date the only genetic mutations described in BWS are in the CDKN1C gene. In order to screen for other genetic predispositions to BWS, the conserved sequences between human and mouse differentially methylated regions (DMRs) of the IGF2 gene were analyzed for variants. Four single nucleotide polymorphisms (SNPs) were found in DMR0 (T123C, G358A, T382G and A402G) which occurred in three out of 16 possible haplotypes: TGTA, CATG and CAGA. DNA samples from a cohort of sporadic BWS patients and healthy controls were genotyped for the DMR0 SNPs. There was a significant increase in the frequency of the CAGA haplotype and a significant decrease in the frequency of the CATG haplotype in the patient cohort compared to controls. These associations were still significant in a BWS subgroup with KvDMR1 LOM, suggesting that the G allele at T382G SNP (CAGA haplotype) is associated with LOM at KvDMR1. This indicates either a genetic predisposition to LOM or interactions between genotype and epigenotype that impinge on the disease phenotype.

INTRODUCTION

Beckwith–Wiedemann syndrome (BWS; MIM 130650) is a congenital overgrowth syndrome, characterized by prenatal and postnatal overgrowth, macroglossia and anterior abdominal wall defects. Additionally, variable features include organomegaly, neonatal hypoglycaemia, hemihypertrophy, urogenital abnormalities and, in about 5% of children, embryonal tumours (most frequently Wilms’ tumour). The genetics of BWS are complex, but involve mutation or altered expression of several closely linked genes associated with cell cycle and growth control in the imprinted 11p15.5 chromosomal region. Genomic imprinting is an epigenetic modification that causes genes to be expressed according to their parent-of-origin. Specific imprinting genes implicated in the aetiology of BWS include the paternally expressed IGF2, KCNQ1OT1 (LIT1) genes and the maternally expressed H19, CDKN1C (P57KIP2) and KCNQ1 genes. Three major BWS subgroups have been identified, familial (15%), sporadic (83%) and those with chromosomal abnormalities (2%) (reviewed in 1–5). Within these subgroups there is a large spectrum of genetic and epigenetic abnormalities.
Among familial cases, about 40% carry mutations of CDKNIC (6–9). The majority of BWS cases are sporadic and these can be subdivided according to their molecular pathology. Thus CDKNIC mutations occur in ~5% and uniparental disomy accounts for another 10–20% of sporadic BWS cases. Up to 60% of sporadic cases are due to epigenetic modifications. These epimutations occur in the two imprinting centres, IC1 and IC2. The first imprinting centre, IC1 is a differentially methylated region (DMR) about 5 kb upstream of H19 and has been shown to have a boundary function regulated by the zinc finger transcription factor, CTCF. The boundary is methylation sensitive, such that when CTCF binds to the unmethylated maternal allele, the IGF2 promoters do not have access to enhancers downstream of H19. Methylation on the paternal allele prevents CTCF from binding, thus permitting interaction between the IGF2 promoters and the enhancers (10,11). About 5–10% of sporadic BWS cases have hypermethylation of the H19 DMR and, in these cases, IGF2 shows loss of imprinting (LOI) and biallelic expression (12). In addition, patients with maternal 1p15.5 chromosome rearrangements have biallelic IGF2 expression, but there are also patients with IGF2 LOI without evidence of IC1 or IC2 mutations (13–16).

The second imprinting centre, IC2, is a DMR located in intron 10 of the KCNQ1 gene and is known as KvDMR1. The unmethylated paternal allele permits transcription of the antisense transcript KCNQ1OT1 (also known as LIT1) and silencing of the KCNQ1 and CDKN1C genes. Maternal methylation at the KvDMR1 is thought to prevent transcription of the KCNQ1OT1 gene and enable expression of KCNQ1 and CDKN1C (14,17–20). It has been proposed that antisense transcripts regulate overlapping genes by promoter occlusion or by competing with these genes for regulatory elements (21). However, current thinking is that RNAi machinery has a role in imprinted gene silencing (22–25). Recent evidence suggests that the KvDMR1 has insulator activity (26–28). Loss of methylation (LOM) at the KvDMR1 is seen in up to 50% of sporadic BWS and in these cases there is KCNQ1OT1 LOI and reduced CDKN1C expression (17).

While it is generally thought that the IC1 and IC2 regions are mechanistically independent (19,29), there are some BWS patients with KvDMR1 LOM that also have IGF2 LOI (1,14,30). In addition, mouse knockouts of Cdkn1c do have abdominal wall defects but are not overgrown; Cdkn1c deficiency needs to be combined with Igf2 over-expression to obtain most BWS symptoms in mouse models (31). Therefore the question of whether there are mechanistic or phenotypic interactions between the IC1 and IC2 regions remains open.

Three DMRs have been identified in the mouse IGF2 gene. These are DMR0 and DMR1 upstream of promoter 3, and DMR2 situated in exon 6 (32,33). In humans, the regions homologous to mouse DMR0 and DMR1 show methylation on either allele. The graph is a dot plot generated by comparing the mouse Igf2 gene, indicated by bars below the genes. Bars with the upper section shaded indicate methylation on the maternal allele, while shading on the lower section indicates methylation on the paternal allele. DMR1 in human is not methylated on either allele. The graph is a dot plot generated by comparing the mouse Igf2 sequence accession number MM71085 with the human IGF2 sequence containing the regions homologous to mouse DMR0 and DMR1.

Figure 1. Differentially methylated regions at the human and mouse IGF2 locus. The mouse and human IGF2 genes are depicted with translated exons (clear boxes) and untranslated exons (dark boxes). Promoters are indicated by arrows. There are three differentially methylated regions (DMRs) in the mouse Igf2 gene, indicated by bars below the genes. Bars with the upper section shaded indicate methylation on the maternal allele, while shading on the lower section indicates methylation on the paternal allele. DMR1 in human is not methylated on either allele. The graph is a dot plot generated by comparing the mouse Igf2 sequence accession number MM71085 with the human IGF2 sequence containing the regions homologous to mouse DMR0 and DMR1.

(36,37), an activator at DMR2 (38) and possibly a promoter region for a placental specific transcript (33,39).

The IGF2 gene encodes a fetal growth factor and is a candidate gene for BWS, since overgrowth in BWS is restricted to those tissues in which IGF2 is expressed. In mice, over-expression of IGF2 results in most of the symptoms of BWS, including prenatal overgrowth, polyhydramnios, fetal and neonatal lethality, disproportionate organ overgrowth and macroglossia (40). Associations between single nucleotide polymorphisms (SNPs) in the IGF2 gene and body mass index in adult males have been described (41). Inter-individual variability in allele specific gene expression (or epigenetic heterogeneity) has been described in some imprinted genes and familial studies suggest there may be a genetic predisposition to epigenetic mutation (42,43). It has also recently been reported that IGF2 loss of imprinting (LOI) in healthy individuals can be a predictive marker for colorectal cancer (44).

Single nucleotide polymorphisms (SNPs) provide an invaluable tool to uncover the basis of multigenic human diseases. The observation of SNPs in specific locations in the genome in
association with observable traits has led to the identification of many candidate genes, as well as key regulatory elements in multigenic diseases. Here we describe SNP analyses in the differentially methylated regions of the IGF2 gene and show that four SNPs in the DMR0 region are in linkage disequilibrium such that only three haplotypes are present in humans. Haplotype analyses in a control population and a cohort of BWS patients uncovered an association with SNPs at the IGF2 DMR0 and BWS, which persisted in a subgroup of patients with loss of methylation in the KvDMR1.

RESULTS

Genetic variation in conserved sequences of differentially methylated regions of the IGF2 gene

Regulatory regions important for imprinted gene expression are likely to be conserved. Figure 1 depicts the three differentially methylated regions of the mouse and corresponding regions in the human IGF2 gene. Previous sequencing of the intragenic IGF2 DMR2 in a cohort of BWS and controls showed no sequence variation in the DMR2, apart from the known polymorphic Apa1 site (45). Dot plot sequence alignment between human and mouse IGF2 was used to identify sequences within DMR1 and DMR0 with >65% homology (Fig. 1B). No variation was found in the DMR1 after 20 DNA samples from unrelated individuals were sequenced (data not shown), suggesting that there are no frequent sequence variations present in this region. By contrast, a high degree of variation was detected in a 626 bp DMR0 homology region. Human DMR0 is located between exons 2 and 3 and has >65% sequence conservation between mouse and human in the 600 bp immediately adjacent to exon 2. Four SNPs were detected in this region. These were at nucleotide positions 123, 358, 382 and 402 in the GenBank sequence Y13633. The SNPs were designated T123C, G358A, T382G and A402G to indicate the variant alleles at each nucleotide position. G358A is a known Msp1 polymorphism.

Linkage disequilibrium between SNPs in IGF2 DMR0

A total of 120 DNA samples from unrelated individuals [40 random human control DNA samples obtained from the European collection of cell cultures (ECACC), plus 20 normal Caucasian laboratory controls and 60 placental DNA samples from a collection of Caucasian families] were genotyped for the SNPs in DMR0. Allele frequencies at each SNP were T123C (T = 0.34, C = 0.66), G358A (G = 0.33, A = 0.67), T382G (T = 0.23, G = 0.77) and A402G (A = 0.46, G = 0.54). The genotype frequencies at each SNP were T123C (TT = 0.15, TC = 0.39, CC = 0.46), G358A (GG = 0.15, GA = 0.37, AA = 0.49), T382G (TT = 0.63, TG = 0.28, GG = 0.09) and A402G (AA = 0.38, AG = 0.31, GG = 0.31). These genotype frequencies did not differ significantly from frequencies expected under Hardy–Weinberg equilibrium with the exception of A402G, which displayed a high degree of homozygosity for both the G and A alleles ($P = 0.0003$).

Haplotype analyses were carried out on 32 of the heterozygous patient samples to confirm that no further haplotypes were present in the BWS population. Haplotypes were assigned to the heterozygous individuals on the basis of the linkage disequilibrium observed in the control population and the allele specific sequencing. There was a significantly altered distribution of TGTA, CATG and CAGA haplotypes ($P = 0.007$) in the BWS patients compared to the controls, such that the ratio of CAGA and CATG haplotypes is different in the BWS population, with CAGA being increased and CATG being reduced. The incidence of the TGTA haplotype was not changed. These results indicate that the T382G SNP (CAGA haplotype) is associated with our cohort of BWS patients.

DMR0 haplotype analyses in BWS patients

A cohort of 73 sporadic BWS patients was genotyped for the DMR0 haplotypes. Table 1 gives the individual SNP and haplotype frequencies (total chromosomes: controls, $n = 236$, and patients, $n = 146$). There was no significant difference in the frequency of the T123C and G358A SNPs between our controls and cohort of BWS patients. There was a significant increase in the frequency of the A allele at A402G ($P = 0.008$) and a significant increase in the frequency of the G allele at T382G ($P = 0.006$) in the patient cohort compared to controls. Twenty five patients were homozygous for all four alleles, with again the third three haplotypes TGTA ($n = 8$), CATG ($n = 7$) and CAGA ($n = 10$). As seen in the controls, the heterozygous patients were heterozygous for the first two plus either the third or fourth SNP or heterozygous at the last two SNPs. Allele specific sequencing was carried out on 32 of the heterozygous patient samples to confirm that no further haplotypes were present in the BWS population. Haplotypes were assigned to the heterozygous individuals on the basis of the linkage disequilibrium observed in the control population and the allele specific sequencing. There was a significantly altered distribution of TGTA, CATG and CAGA haplotypes ($P = 0.007$) in the BWS patients compared to the controls, such that the ratio of CAGA and CATG haplotypes is different in the BWS population, with CAGA being increased and CATG being reduced. The incidence of the TGTA haplotype was not changed. These results indicate that the T382G SNP (CAGA haplotype) is associated with our cohort of BWS patients.

Our BWS cohort has been molecularly characterized and consisted of two subgroups: 44 patients with loss of methylation (LOM) at the KvDMR1 and 29 patients with mosaic uniparental disomy (UPD) for 11p15.5. Subset analyses
were carried out to see if we could detect a significant association between the CAGA haplotype and patients with either UPD or KvDMR1 LOM.

The KvDMR1 LOM cohort (88 chromosomes) compared to the controls still showed a significant difference in haplotype distribution ($P = 0.013$), with the T382G variant defining the CAGA haplotype still increased ($P = 0.009$) and the A402G variant defining the CATG haplotype reduced ($P = 0.015$) (Table 2). These results show that the CAGA haplotype is associated with BWS patients that have loss of methylation at KvDMR1. The haplotype distribution in the KvDMR1 LOM cases was TGTA (30), CATG (26) and CAGA (32). There were 12 patients homozygous for the haplotypes TGTA ($n = 5$), CATG ($n = 2$) and CAGA ($n = 5$). The homozygous frequencies were as expected under Hardy–Weinberg equilibrium for the observed allele frequencies. Only six cases in our cohort of KvDMR1 LOM patients were informative for IGF2 LOI, five of these cases had the CAGA haplotype.

In the UPD group, the constitutive and disomic haplotypes could be distinguished (see Materials and Methods). A homozygous result in the UPDs could potentially reflect a constitutive homozygote or a high percentage of disomy. Table 3 shows that 10 out of the 13 homozygous cases in this cohort had a high average ratio of UPD to normal cells. Five cases showed equal ratios for heterozygous SNPs and these cases all had a UPD to normal ratio of $\leq 4.5$, suggesting that we could be sampling either the disomic cell line (UPD ratio $\geq 4.5$) or the normal cell line (UPD $\leq 4.5$) in these cases. We decided to exclude the homozygous cases with UPD ratios $>5$ so as to assign unequivocal constitutive haplotypes to the UPD patients. Thus constitutive haplotype distribution in the UPD patients was TGTA (16), CATG (11) and CAGA (13). This distribution did not differ significantly from the control ($P = 0.11$) and although the CATG haplotype was significantly reduced compared to controls ($P = 0.04$), the CAGA haplotype was not significantly increased ($P = 0.15$) (Table 2). The removal of nine homozygous UPD patients from the total patient population resulted in the BWS haplotype distribution totals being TGTA (46), CATG (37) and CAGA (45), which is still significantly different from the control population ($P = 0.005$), and with the CATG haplotype significantly reduced ($P = 0.004$) and the CAGA haplotype significantly increased ($P = 0.007$).

In 24 of the UPD samples it was possible to distinguish unequivocally the disomic haplotype, these included the 13 homozygous cases plus 11 heterozygous cases with one haplotype more prevalent than the other. The haplotype distribution in the disomic cells was TGTA (16), CATG (18) and CAGA (14). This distribution was not significantly different to the controls ($P = 0.52$) and interestingly did not enrich for the CAGA haplotype as the expressing haplotype.

Thus sub-analyses of the normal and disomic haplotypes in UPD patients do not show a significant association with any DMR0 haplotype. This is still a small sample and it is still possible that significant differences in haplotype ratios could be detected when more patients with UPD are added to the study.

To examine whether the DMR0 haplotypes had any effect on IGF2 transcription, we cloned the three different DMR0 haplotypes into a P3 luciferase construct (38), and assayed the effect that DMR0 sequences had on transcription. Transient transcription assays indicated slight but not significant enhancing activity for the TGTA haplotype, while the CAGA and CATG haplotypes had no effect on transcription (data not shown). Quantitative PCR was used to measure IGF2 mRNA
levels in seven control placental samples that were homozygous for the DMR0 haplotypes. At this level of analysis, high or low expression levels could not be associated with any particular haplotype (data not shown). However, this does not exclude effects at other developmental stages or tissues which are not easily accessible for human studies. If allelic variation at the DMR0 is not associated with variation in IGF2 transcription, it still does not exclude an interaction between DMR0 haplotypes, IGF2 LOI and other epigenetic changes.

**DISCUSSION**

Epigenetic alterations in IGF2 imprinting have been implicated in many BWS patients, but often the underlying causes are still unknown. Here we have investigated the possibility that genetic variants in the IGF2 gene might be associated with BWS. We discovered new SNPs in one of the differentially methylated regions in IGF2 which is conserved between mouse and human DMR0. Our cohort of sporadic BWS patients consisted of those with UPD and those with loss of methylation at KvDMR1. Together these groups constitute ~70–80% of sporadic BWS cases (1,3). We found a highly significant association between a specific IGF2 DMR0 haplotype (CAGA) and our BWS patients ($P = 0.007$). This association is not significant in the subgroup with UPD ($P = 0.15$), but is still highly significant in the subgroup of patients with IC2 epimutations ($P = 0.013$). We cannot exclude that there are also associations between the CAGA haplotype and sporadic BWS patients with mutations within CDKN1C or with H19 imprinting lesions. However, these rare subgroups of BWS make up <10% of our sporadic patients, precluding an analysis of association because of small numbers. Notwithstanding, this is the first association discovered of genetic variants in IGF2 with BWS, and suggests that such variants interact with loss of methylation at IC2 to cause BWS.

DMR0 in Igf2 was first discovered in the mouse as a maternally methylated DMR overlapping the placental-specific promoter P0 (33). LOI of P0 expression is associated with LOI at DMR0 (33). Evidence for the human IGF2 DMR0 having a role in the regulation of IGF2 imprinting comes from observations of hypomethylation of this region in Wilms’ tumour and colorectal cancer and also in some normal healthy individuals that show relaxed imprinting at Igf2 (34,35,42–44). It has been proposed that the IGF2 DMR0 is a methylation dependent silencer element that may be regulating IGF2 expression independently from the H19 DMR and that the regulation and maintenance of imprinting IGF2 expression does not depend entirely on the imprinting centres within the cluster (35). Biallelic IGF2 expression on its own is not enough to cause BWS or cancer and it is estimated that 10% of normal healthy individuals have a constitutive loss of IGF2 imprinting. Therefore it is possible that the variants in DMR0 are associated with altered states of IGF2 expression or imprinting. Our preliminary results on determining expression levels of IGF2 in placenta with the different DMR0 haplotypes do not allow us to reject this hypothesis, since a comprehensive analysis of different IGF2 transcripts in different tissues could not be carried out. Equally, there were not enough patients with LOI of IGF2 to establish or refute a connection with the DMR0 haplotypes. Nevertheless, an association of DMR0 haplotypes with IGF2 levels would make biological sense, particularly because we still find this association in patients with loss of methylation at KvDMR1. Thus it is possible that LOI at KvDMR1 on its own does not result in the full spectrum of BWS symptoms but must be associated with particular IGF2 alleles described here for full expression of the phenotype. LOI at KvDMR1 leads to partial silencing of CDKN1C (17,18). BWS patients with CDKN1C mutations may also be associated with particular IGF2 haplotypes, but we could not test this due to the small number of sporadic cases with such...
mutations. In the mouse knockout of Cdkn1c there is placental hyperplasia and abdominal wall defects, but not pre- or postnatal overgrowth (46,47). Indeed this knockout needs to be combined with over-expression of Igf2 in order to obtain the full spectrum of BWS phenotypes (31).

An alternative explanation of the association between DMR0 SNPs and LOM at KvDMR1 is that this reflects a genetic predisposition to epimutation at IC2, as well as at IGF2. A genetic component to epimutation is evident in mice where strain dependent relaxation of imprinting of Kcnq1 was observed (48). In humans, a number of familial studies where kindred share loss of imprinting has been reported (43,49). Familial clustering of IGF2 LOI together with haplotype analyses of these families have indicated that the loss of imprinting is likely to be due to trans effects (43). However, in at least one family with heritable LOI of IGF2, the same maternal allele segregated with LOI, indicating a cis acting variant which predisposes individuals to LOI in this family. In our dataset it was not possible to rigorously investigate whether DMR0 haplotype was associated with the presence or absence of IGF2 LOI in BWS patients with KvDMR1 LOM as appropriate samples were only available from a subset of patients. However, we have found in preliminary studies some cases with KvDMR1 LOM and IGF2 LOI in which DMR0 was hypomethylated, suggesting the possibility of an association between DMR0 genotypes and epimutations at DMR0 leading to LOI. Clearly it would be of interest to know whether DMR0 haplotype was associated with the presence of DMR0 hypomethylation and IGF2 LOI in normal individuals and colorectal cancer cases.

A bioinformatics search to detect transcription factor binding sequences overlapping the SNP positions indicated that the T382G SNP overlapped the early developmental transcription factors forhead box HNF3B(FOXA2), GATA-4 and LM02 binding sites. The CAGA haplotype would abolish these sites. HNF3B and GATA-4 have been shown to open compacted chromatin in vitro through specific transcription factor-histone interactions (50). Thus IGF2 alleles with the CAGA haplotype may tend to be more compacted and less easily remodelled making it resistant to imprinting erasure, reprogramming or maintenance of imprinting. Variation in susceptibility in chromatin remodelling may also explain why we could not show that the CAGA haplotype directly influences IGF2 transcripion levels in normal placental tissue in a normal epigenetic environment.

Could DMR0 SNP variants also constitute a genetic predisposition to LOM at KvDMR1 which is about 500 kb away? It is generally recognized that many genetic associations may derive from linked markers, necessitating exploration of the haplotype structure at a given locus. The discovery of allelic variation in the DMR0 and the association of particular haplotypes with LOM at the distant KvDMR1 could imply regulatory elements that influence long range chromatin structures including methylation at the KvDMR1. Evidence for long range interactions between genes at this locus comes from transgenic studies in mice where expression of the Cdkn1c gene was shown to depend on enhancers as far as 250 kb downstream (51). In this study a BAC extending 250 kb downstream of the Cdkn1c gene and containing the Kcnq1ot1 (lit1) gene could fully recapitulate embryonic Cdkn1c expression. However, differential methylation of Cdkn1c and the KvDMR1 could not be established after germline transmission of this BAC, suggesting that additional cis-regulatory elements outside the 250 kb region are required for methylation (51). An alternative explanation is that T382G, the defining SNP of the CAGA haplotype may be in linkage disequilibrium with another regulatory SNP closer to the Kcnq1 and its antisense KCNQ1OT1 (LIT1) gene. LD is more often a property of the chromosomal region than a sole reflection of the physical distances between these markers. The low amount of haplotype diversity in this region to the degree that only three haplotypes could be detected, infers that the DMR0 is situated within a region of low recombination (i.e. a haplotype block). The likelihood of the T382 SNP being in non-random association with a SNP closer to the KCNQ1 locus would depend on the extent of the haplotype block, which can be anything from 2 to 800 kb. Extensive analyses of the KCNQ1OT1 (LIT1) region for polymorphism was not carried out, however in 14 cases where there was information available on a SNP in the KCNQ1OT1 [SNP2 in (19)], this was independent of DMR0 SNPs.

Further work is required to identify the triggers for LOM at KvDMR1 and the significance of the role that the DMR0 haplotypes play in this process. Nevertheless, this work is the first to show an association between genetic and epigenetic variants in an imprinting cluster. Interactions between genetic

<table>
<thead>
<tr>
<th>Table 3. Genotype and mosaic UPD ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMR0 genotype ratio</td>
</tr>
<tr>
<td>Homozygous</td>
</tr>
<tr>
<td>Heterozygous</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

*UPD normal ratio was determined by markers on chromosome 11. Up to five markers (see Materials and Methods) were tested in each case and the average ratio (paternal allele: maternal allele) was determined. The value ‘5’ indicates that the ratio was large, (i.e. a very low value for the maternal peak).

Genotype ratios were determined by examining peaks at the individual SNPs on sequence chromatograms. Homozygous genotypes showed no variation. Heterozygous genotypes showed varying peaks proportional to the mosaic UPD ratio.
and epigenetic factors could be of importance not only in imprinting syndromes, but also particularly in cancer and in multifactorial common diseases (such as for example diabetes or schizophrenia) in which genetic factors have been difficult to pin down. Perhaps interactions with epigenetic factors have made pure genetic analyses difficult; this situation may improve in the future with the possibility of combined genomic and epigenomic analysis.

MATERIALS AND METHODS

Genomic DNA

Genomic DNA was isolated from either fibroblast or lymphoblast cell lines, placentae or the peripheral blood of patients and controls.

Normal control DNA samples

Forty control DNA samples were obtained from the EACC and consisted of third generation healthy Caucasian individuals from the UK and Ireland. (This is a collection of 500 lymphoblastoid cell lines derived from randomly selected Caucasian blood donors whose parents and grandparents were born in the UK or Ireland.) A further 20 DNA samples were from the peripheral blood samples taken from parents of children presenting for cystic fibrosis testing at the West Midlands Regional Genetics Centre at Birmingham Women’s Hospital. Sixty DNA samples were extracted from placentae from a collection of families having a baby at Queen Charlotte’s and Chelsea Hospital. These samples were a subset of a population of 250 white-background trios collected from families who have given written consent for parental blood and placental collection. The collection of samples has been approved by the joint Hammersmith Hospitals Trust and Imperial College Ethics Committee.

Patient samples

A cohort of 44 sporadic BWS patients were selected on the basis of loss of methylation at the KvDMR1 determined as described previously (14,52) and by a bisulphite assay. The second cohort of 30 UPD patients were characterized on the basis of informative markers at the 11p15.5 locus (D11S2071, D11S1984, D11S4046, D11S1318, D11S4088, D11S4177, D11S4146). All the BWS patients were homozygous for the CAGA (two samples), CATG (two samples) haplotypes. Real time quantitative PCR assays

cDNA samples were extracted from seven placentae which were homozygous for the CAGA (two samples), CATG (two samples) and TGTA (three samples) haplotypes. Real time PCR was carried out using Applied Biosystems ABI Prism 7700, and SYBR Green master mix (Applied Biosystems) in a 25 μl reaction and standard cycling conditions of 50°C, 2 min, 95°C, 10 min followed by 40 cycles of 95°C,15 s and 60°C, 1 min. 1. IGF2 primers were forward: GAA GGT GAA ATT C. These primers were used at concentrations of 250 nM, 200 nM, 100 nM, and 50 nM.

Luciferase assays

Reporter constructs were made using TGTA, CAGA and CATG sequences obtained from the cloned PCR products from the allele specific sequencing experiments. The DMR0 haplotypes were cloned upstream to the mouse Igf2 promoter 3 in a luciferase construct and the luciferase assays were performed in triplicate in a HEK 392 cell line as previously described (38).

Sequencing analysis

PCR products were cloned into PCR 2.1 cloning vectors (Invitrogen) prior to sequencing and up to six independent bacterial cultures sequenced per patient or control sample. The PCR product was purified using QIAquick PCR Purification Kit (Qiagen) to remove primers and dNTPs and then sequenced by an external sequencing service (Lark Technologies Inc., GRI-Genomics). The sequencing reaction could detect all four SNPs in one reaction, and since both forward and reverse sequencing was carried out on all the samples, there was no ambiguity using this method. Sequences were analyzed using Chromas 1.45 software (http://www.technelysium.com.au) and the chromatogram of each sequence was examined visually in order to assess whether a SNP was homozygous or heterozygous.

In the UPD patients group, the constitutive and disomic haplotypes could be established by examining the amplitude of the chromatogram peaks at the various SNPs. Thus where a normal non-mosaic heterozygote would have equal amplitudes for both peaks, a UPD patient that is mosaic for disomic and constitutive heterozygous cell lines would show a larger peak for the disomic allele and this peak would be consistent at all four SNPs. The amplitude depends on the percentage of mosaicism and a homozygous result could potentially reflect a constitutive homozygote or a high percentage of disomy.

Allele specific sequencing to determine haplotypes

PCR products were cloned into PCR 2.1 cloning vectors (Invitrogen) prior to sequencing and up to six independent bacterial cultures sequenced per patient or control sample.

Quantitative real time PCR assays

cDNA samples were extracted from seven placentae which were homozygous for the CAGA (two samples), CATG (two samples) and TGTA (three samples) haplotypes. Real time PCR was carried out using Applied Biosystems ABI Prism 7700, and SYBR Green master mix (Applied Biosystems) in a 25 μl reaction and standard cycling conditions of 50°C, 2 min, 95°C, 10 min followed by 40 cycles of 95°C,15 s and 60°C, 1 min. 1. IGF2 primers were forward: GAA GGT GAA ATT C. These primers were used at concentrations of 250 nM, 200 nM, 100 nM, and 50 nM.

Luciferase assays

Reporter constructs were made using TGTA, CAGA and CATG sequences obtained from the cloned PCR products from the allele specific sequencing experiments. The DMR0 haplotypes were cloned upstream to the mouse Igf2 promoter 3 in a luciferase construct and the luciferase assays were performed in triplicate in a HEK 392 cell line as previously described (38).

Sequencing analysis

PCR products were cloned into PCR 2.1 cloning vectors (Invitrogen) prior to sequencing and up to six independent bacterial cultures sequenced per patient or control sample. The PCR product was purified using QIAquick PCR Purification Kit (Qiagen) to remove primers and dNTPs and then sequenced by an external sequencing service (Lark Technologies Inc., GRI-Genomics). The sequencing reaction could detect all four SNPs in one reaction, and since both forward and reverse sequencing was carried out on all the samples, there was no ambiguity using this method. Sequences were analyzed using Chromas 1.45 software (http://www.technelysium.com.au) and the chromatogram of each sequence was examined visually in order to assess whether a SNP was homozygous or heterozygous.

In the UPD patients group, the constitutive and disomic haplotypes could be established by examining the amplitude of the chromatogram peaks at the various SNPs. Thus where a normal non-mosaic heterozygote would have equal amplitudes for both peaks, a UPD patient that is mosaic for disomic and constitutive heterozygous cell lines would show a larger peak for the disomic allele and this peak would be consistent at all four SNPs. The amplitude depends on the percentage of mosaicism and a homozygous result could potentially reflect a constitutive homozygote or a high percentage of disomy.
Statistical analysis

Chi square tests and Fisher's exact analysis for 2 × 2 contingency tables (conditional on marginal totals) were set up for individual SNP frequencies. Odds ratios and confidence intervals were calculated from the same 2 × 2 contingency tables. Similarly, the overall difference between haplotype distributions in case and controls was studied using 2 × 3 contingency tables. For linkage disequilibrium studies the method of Taillon-Miller (53) was used to make intermarker LD (pairwise) comparisons and to determine the normalized coefficient of disequilibrium D'. LD was calculated using individuals homozygous for all four SNPs (92 individuals, n = 184 chromosomes). Composite linkage disequilibrium was analyzed using GDA software (available at http://lewis.eeb.uconn.edu/lewishome).

ACKNOWLEDGEMENTS

We are grateful to Bryan Barrat, Clive Petry and Eurot Walter for helpful statistical discussions; Sayeda Abu-Amero for technical assistance; Fatima Santos, Andrea Riccio and Gavin Kelsey for reading the manuscripts and Dirk Prawitt for helpful discussion. Work on this manuscript was funded by the CR-UK and the MRC.

REFERENCES


